THE TRANSPORT DETECTIVES

A diverse team of data scientists from LTA is poring over data to provide insights on travel behaviour, traffic conditions and even trends relating to where people live and work, with a view to improving commuter experiences.

Armed with smart analytics tools, a team of data scientists is cracking codes. Behind screens, they scrutinise numbers, algorithms and charts, trying to find patterns in the clutter.

With statistics from EZ-link card transactions, they map out hotspots across Singapore indicating the volume of commuters at transportation nodes.

With rainfall information from the National Environment Agency (NEA), the team assesses if the supply of taxis is adversely affected due to wet weather.

With drone footages captured in the sky, the scientists analyse traffic conditions by detecting the number of vehicles at busy road junctions.

Their mission: Using data to make commuting easier.

HOW LTA BEGAN TO HARNESS THE POWER OF BIG DATA

Not many people are aware that LTA is a data-driven agency with a special team of data scientists and analysts.

The data science division began in 2013. But the genesis of the team goes back to the mid-2000s.

"In 2008, we had a headache. We had a huge volume of data – 11 million records of information every day from our EZ-link transactions when we tap in and tap out," said Jonah Ong (pictured, 2nd from left), who heads the division.

That was when the IT engineers started to think of a way to create value with the amount of information they possessed.

Jonah adds that,

"Coincidentally, the term 'big data' started to become popular all over the world. We looked at ourselves and realised that we were already doing it."

In 2010, LTA launched the Planning for Land Transport Network (PLANET) system. It conducted advanced data analytics on travel patterns of Singaporeans whether by car, bus or train – with the aim to enhance data-driven decision-making.

Now, the team is making use of better tools, as well as more abundant and accessible data, to do things today that they couldn't just a few years ago.

A DIVERSE TEAM OF "SKUNKWORKS"

The data science team comprises 12 members who have different specialisations. There are geographers, mathematicians, economists, statisticians and even a chemical engineer among them.

"We wanted to hire everyone to be unique on their own and bring something different to the table."

JONAH ONG, DEPUTY DIRECTOR, DATA SCIENCE

"So, like a patchwork, they can look at a picture from different angles to solve a problem."

"People refer to us as 'skunkworks'," he added, referring to the term used to describe a team put together informally to solve a problem.

"Data looks boring at first. You see rows and columns of numbers and tags. But what's key is to be observant and see the clues in the data, and sniff out what people cannot see."

This division anticipates trends and responds quickly to issues.

"Data science is agile by nature. It is highly iterative and unpredictable," said Jonah who runs the team like a start-up.

"So the way we run the team has to be nimble, so we can respond to business needs faster."

Project Spotlight

IS IT HARDER TO GET A TAXI WHEN IT RAINS?

We all feel that cabs disappear when you need them the most, especially during wet weather. Data analyst Rachel Lim (pictured right) wanted to check if this was really true.

She used rainfall data from NEA, which is collected at five-minute intervals at 55 rain gauges islandwide, as well as data of taxi locations.

"We found that during prolonged periods of rainfall, taxi availability tends to decrease marginally in some parts of Singapore," noted Rachel, who is one of the youngest persons on the team.

"But the effect is not as pronounced as we think. There are other factors like time of day, and peak periods."

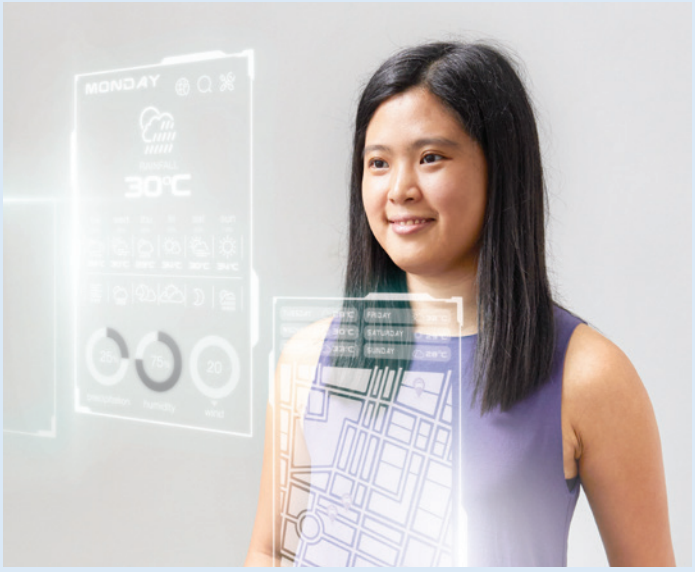
Beyond that, analysing rainfall data has other applications too – including the correlation to bus speeds, which may have implications on bus planning.

Rachel read geography for her undergraduate degree, as she is interested in city planning and urban design.

"I wanted to join the public service because I wanted to contribute to something," she said. "LTA stood out because during the interview process, it wasn't simply about what I had to offer. It felt like an organisation that cared, and that there were opportunities for growth."

She later pursued a Masters in urban data science, saying,

"I felt that geography theories and planning alone were insufficient. So I thought the two would complement each other well."



CAN WE INFER WHERE PEOPLE STAY AND WORK USING THE GINI CONCEPT?

EZ-link card transactions reveal two common locations that we often travel to – home and office.

But how does data interpret which is home or and and which is the office?

To do so, data analyst Cheo Rui Ming (pictured above) uses the economics concept of Gini.

Traditionally, the Gini coefficient is used to measure income inequality. It ranges from a value of 0 to 1, where 0 is complete equality, and 1 is complete inequality.

"When you are at home, the distribution of the hours spent is more varied. No one can be sure how long you spend at home," said Rui Ming. "Whereas at work, time spent there is more uniform."

Translating this to the Gini concept, a value close to 0 means time spent at a location is more consistent – or equally distributed – than a value closer to 1.

This means LTA would be able to zoom in on residential and work hotspots to better size up plans for land transport infrastructure.

Rui Ming, an economics graduate, had gone on a 12-week accelerated private course on data science, and met Jonah who was presenting on LTA's work.

"I found out that the work that LTA did was diverse," he said. "I thought joining the team would allow me to explore and expose me to different domains."

"Data science was an opportunity for me to take my economics training a step further and to sharpen my skills."

CAN WE ANALYSE TRAFFIC CONDITIONS ON THE ROAD USING DRONE FOOTAGE?

Fixed cameras are usually used to monitor traffic conditions along expressways and at major road junctions. But there are locations with limitations where cameras are not available.

Instead, the data science team uncovered another way to detect and count vehicles – looking at video footage captured by flying drones. This is useful when implementing a traffic scheme change, such as changing the timing of traffic lights.

The analytical approach can detect and count vehicles from a static or moving image, and can even differentiate between cars, buses, motorcycles and pedestrians, said senior data scientist Leong Wei Deng (pictured below), the brain behind the project.

In time to come, LTA would be able to reduce the manpower needed for site visits to manually count the number of vehicles and assess traffic conditions. Wei Deng's interest in transport planning was piqued after experiencing two train breakdowns, and he now uses drone video footage to analyse traffic conditions.

"I started to realise that transport is something we take for granted. You don't realise the impact transport has in our lives until something big happens."